Article

Exploring Impact of Age and Gender on Sentiment Analysis Using Machine Learning

Sudhanshu Kumar [1,*], **Monika Gahalawat** [1], **Partha Pratim Roy** [1] **and Debi Prosad Dogra** [2] **and Byung-Gyu Kim** [3,*]

[1] Department of Computer Science & Engineering, Indian Institute of Technology Roorkee, Uttarakhand 247667, India; monikagahalawat62@gmail.com (M.G.); proy.fcs@iitr.ac.in (P.P.R.)
[2] School of Electrical Sciences, Indian Institute of Technology Bhubaneswar, Odisha 752050, India; dpdogra@iitbbs.ac.in
[3] Department of IT Engineering, Sookmyung Women's University, Seoul 04310, Korea
[*] Correspondence: skumar2@cs.iitr.ac.in (S.K.); bg.kim@ivpl.sookmyung.ac.kr (B.-G.K.)

check for updates

Abstract: Sentiment analysis is a rapidly growing field of research due to the explosive growth in digital information. In the modern world of artificial intelligence, sentiment analysis is one of the essential tools to extract emotion information from massive data. Sentiment analysis is applied to a variety of user data from customer reviews to social network posts. To the best of our knowledge, there is less work on sentiment analysis based on the categorization of users by demographics. Demographics play an important role in deciding the marketing strategies for different products. In this study, we explore the impact of age and gender in sentiment analysis, as this can help e-commerce retailers to market their products based on specific demographics. The dataset is created by collecting reviews on books from Facebook users by asking them to answer a questionnaire containing questions about their preferences in books, along with their age groups and gender information. Next, the paper analyzes the segmented data for sentiments based on each age group and gender. Finally, sentiment analysis is done using different Machine Learning (ML) approaches including maximum entropy, support vector machine, convolutional neural network, and long short term memory to study the impact of age and gender on user reviews. Experiments have been conducted to identify new insights into the effect of age and gender for sentiment analysis.

Keywords: sentiment analysis; social media; machine learning; lexicon

1. Introduction

The growth of the internet has led to a huge influx of data that holds vast and valuable insights about the public opinion. Every internet user who expresses an opinion on the web becomes a part of this information circuit where other users benefit from these public reviews and hence can make an informed decisions. With the data collected (reviews, posts, comments) from different social media platform such as Facebook, Twitter, Amazon, Goodreads, IMDb or blogs, the task of using these reviews to find the polarity of public (positive, negative or neutral) opinion is called Sentiment analysis. Sentiment analysis is generally performed on movie reviews [1,2], restaurant or food reviews [3,4], along with data from microblogs [5,6], providing some useful insights to different organizations to improve business strategies by attracting new customers. The categorization of customers based on age and gender present an important information that can make products more effectively fullfill the demands of different age and gender group persons. This fine-grain information about customers are value-added to enhance the revenue of the company and its reputation in the global market. E-commerce companies want to know the mindset of the customers. For example, females do more

shopping in comparision to male in the E-commerce site and certain portals such as firstcry website are more popular for various products of different age groups including newborn, infant, toddler, etc. Sentiment analysis [7–10] has evolved over the years with different dictionary-based and machine learning techniques implemented to obtain better accuracy. With the advent of deep learning techniques [11–13] in sentiment analysis, prior information has also played a big role in adequately expressing the polarity of opinions. Kahaki et al. [14] proposed an age estimation system based on orthopantomographs images. The orthopantomogram is a dental X-ray of the upper and lower jaw. The geometric mean projection transform was used on Malaysian children dental development dataset with 456 patient's X-ray images to extract and classify the 3rd molar teeth in the orthopantomography images. The proposed system results showed a reliable age estimation method. The importance of age estimation may also be useful for civil, criminal, law enforcement, airport security and for forensic purposes. In a similar study on the same Malaysian children dataset, the automatic age assessment [15] was proposed based on pre-trained deep convolution neural network. The results of this approach concluded that the method can efficiently classify the images with high accuracy and precision.

Li et al. [16] proposed a framework providing an abstract of the opinions using sentiment analysis. The authors have taken into consideration the opinion subjectivity and user credibility in their proposed approach. Lockenhoff et al. [17] analyzed how different age groups express their emotions. The authors found that older adults describe the positive emotions better than their negative emotions as compared to younger adults. Another research by Zimmermann et al. [18] for emotion regulation based on age and gender [19] of the person was able to handle emotions in a better way. The results showed that people in middle adolescence showed the least emotional regulation as compared to other age groups. Even gender differences were also encountered as either under or over estimating a particular emotion. Thus, based on these established psychological differences between people of different age and gender, we aim to find out whether these differences can also be observed in the opinions that the individuals express on on-line platforms.

This study explores the differences in sentiment expressing abilities of different groups and their subsequent impact on the sentiment reviews. This can be extremely helpful for commercial applications as they can focus directly on a particular audience that is more receptive towards their product rather than making a generalized marketing strategy. It will further help in differentiating their brand from other leading companies to provide better customer support. Oh et al. [20] studied the market segmentation on basis of gender and age of users to find the travel potential of different groups based on their incomes and available leisure time. Keshari et al. [21] analyzed the effectiveness of advertising appeals on different gender and age groups based on how the consumers respond to these advertisements. Figure 1 demonstrates the basic flow diagram of the framework used for sentiment analysis, where the data collected from social media is used to extract two datasets on basis of age and gender. The different sentiment analysis approaches have been implemented on this data. The main contributions of the paper are as follows:

1. Exploring the impact of user expression based on age and gender using different feature extraction methods.
2. We create a dataset that contains user reviews along with the user's age and gender information.
3. A detailed analysis on the impact of user expression is presented through extensive experiments.
4. Finally, a comparison with different machine learning and a dictionary-based classifier is also discussed.

The rest of the paper is organized as follows. In Section 2, we discuss the existing research work in sentiment analysis. In Section 3, the methodologies implemented on the dataset have been discussed, along with a comparison of different approaches. Section 4 describes the experimental results with dataset description. Finally, in Section 5 the work has been concluded along with discussion of some future possibilties.
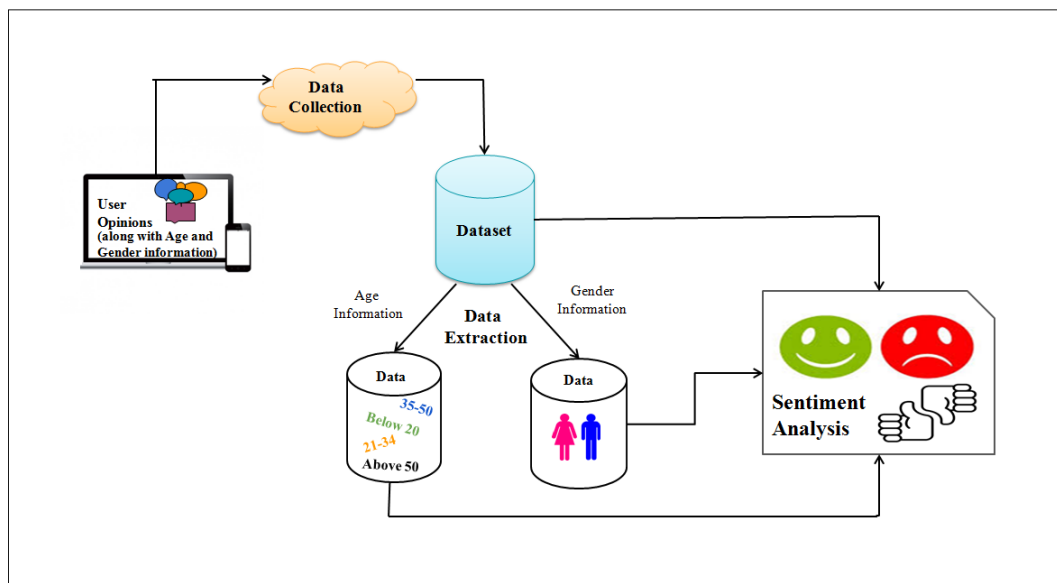
Figure 1. Data points are collected from the social media along with the user's age and gender information. Sentiment analysis is then performed on the newly created data sets.

2. Related Work

In this section, we discuss the recent works of sentiment analysis as researchers try to find a better approach to predict the sentiment polarity. Twitter and Facebook have been the most popular social media platforms as people express their opinion about every topic on these social networking sites, which helps in understanding public sentiment. Appel et al. [22] used twitter sentiment and movie review datasets to implement a hybrid approach based on ambiguity management, semantic rules, and sentiment lexicon. The authors compared this proposed hybrid system results with the standard supervised algorithms such as Naive Bayes (NB) and Maximum Entropy (ME). The proposed system achieves higher precision score and accuracy than the supervised algorithms. Similarly, Zainuddin et al. [23] used a twitter dataset of aspect-based sentiment analysis to perform a fine-grained analysis. They proposed a hybrid approach using a feature selection method that performs better than the standard methods.

Blogs have been a relevant source of data in sentiment analysis with posts containing reviews and comments. Fan et al. [24] analyzed blog text to improve the quality of advertisements in the blogs that were more relevant to the user. To find the blogger's overall emotions towards any particular topic, Kuo et al. [25] create a social opinion graph as generally every blogger is somewhat influenced by its social circle. So their social interactions can be used to find the overall sentiment orientation of the blogger. Li et al. [26] used opinions expressed on the web such as blogs, reviews and comments to design a new technique to further enhance the accuracy of clustering based approaches. This approach is proven to more suitable in finding neutral opinions. The authors [27] proposed a new extraction and opinion mining system based on a type-2 fuzzy ontology called T2FOBOMIE. The proposed system received input from a user, extracts the relevant features from an input query and then converts into to a search query with hotel reviews. The feature opinions, user requirements and hotel information were integrated in a T2FOBOMIE system to achieve high performance.

Apart from using products, movie, restaurants or book reviews for sentiment analysis, researchers have also focused on analyzing sentiment in other languages than English. Pak et al. [28] have proposed a technique that works quite well for other languages as well, though they have not tested their algorithm on multilingual data . The author [29] has implemented a methodology to find sentiment polarity within a multilingual framework and the testing was performed using movie reviews in German language collected from amazon. Similarly, Zhou et al. [30] translated Chinese reviews to English language and then used English language corpus to perform sentiment analysis on these

translated reviews. The authors presented that translated reviews outperform original reviews. Another study on Chinese public figures has been performed in [31] to analyze the opinion polling of public figures.

The analysis of opinions expressed by people from different genders or different age groups should align with their psychological differences, as is illustrated by different research groups. There have been multiple research studies on how different individuals handle different emotions and the way these individuals express their emotions even before the advent of internet. The authors [32] examined gender differences in conducting a study on 400 college students in five age groups from preschoolers to adults. The study aligned with the stereotypes of gender and age emotional expressiveness. Stoner et al. [33] considered people of both genders and in different age groups to study their anger expressing ability. The research showed that young adult group expressed anger more as compared to old adult age group. In this study, the author did not find out much differences on basis of gender in this aspect.

A research by Davis [34] on gender differences in negative emotions showed that boys expressed a greater negative affect as compared to girls when they were disappointed. Brody et al. [35] researched more on gender and emotional expression and showed that gender differences in emotional expressiveness were culturally specific in asian international students. Another study by Kring et al. [36] in which they showed emotional videos to a group of students and reaffirmed that women are generally more expressive than men even in case of experienced emotions. A study by Birditt [37] examined age and gender differences in description of emotional reactions. It contained 185 individuals as 85 males and 100 female aged from 13 to 99 which showed that adolescents and young adults were reported more likely to describe anger and giving more intensive aversive responses as opposed to the male adult group.

3. Methodology

To process the reviews, the steps in the Figure 2 are followed. Firstly, the dataset segregated into two sets on basis of age and gender and then separated into categories based on the specific age and gender. Secondly, each particular data group is divided in training data and testing reviews.
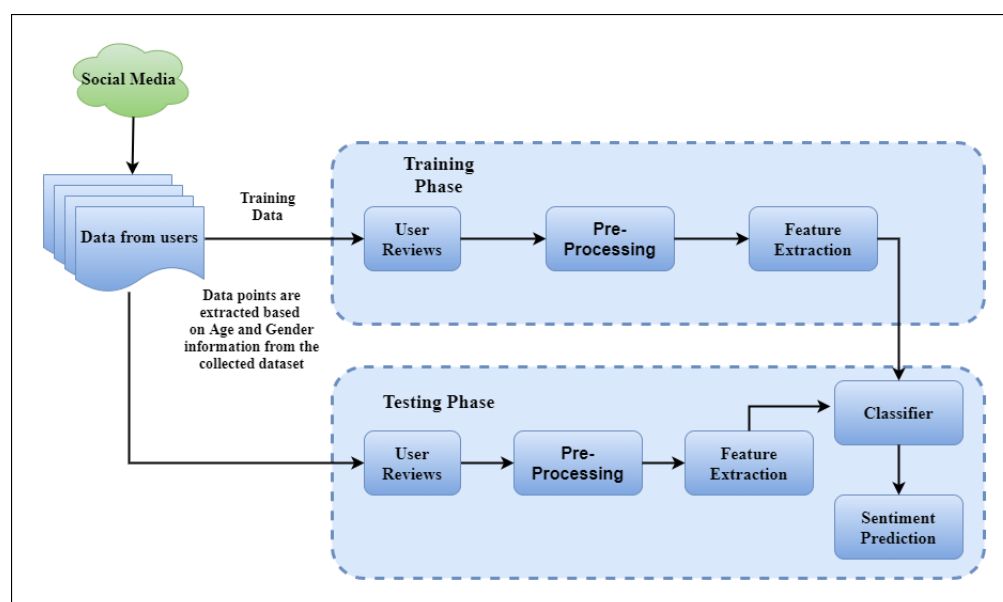

Figure 2. Flow Diagram representing the steps taken for sentiment analysis where the classifier algorithm is implemented at the end of training after the data pre-processing and feature extraction and it used in testing step to produce the final results. The user reviews need to go through pre-processing and feature extraction in the testing phase as well before being passed on to the classifier algorithm.

The reviews [38] are pre-processed to remove the unnecessary information from the reviews that has no effect on the polarity of the sentence. So, we perform data cleansing through the steps as shown in Table 1. Then, the feature extraction steps are performed as explained in Section 3.1.2. Finally, the classifier algorithm predicts the label which when compared to the ground truth gives the accuracy of the classifier. We have collected data regarding people's preference for the books (hard cover, kindle ebooks or audio books) along with their age and gender information. We implement different algorithms for sentiment analysis on each set of data separately and the results are then compared to identify the respective differences between the groups. Also, a dictionary-based approach has been implemented on the collected dataset.

Table 1. Pre-processing steps that have been performed on the user reviews for doing data cleansing and removing uninformative parts that has no effect on the sentiment score of the sentence.

S.No.	Description of Noisy and Uninformative Parts in Reviews
1.	Removing punctuations, numbers and symbols since they do not add any substantial meaning to the sentence that may affect it's sentiment score.
2.	Removing stop words as they make no impact on the sentiment score of the expressed opinion.
3.	Replacing the acronyms of a word with the actual word.
4.	Transforming the text to lowercase.
5.	Replacing emoticons with the sentiment that the emoticon expresses.
6.	Tokenize the review.

3.1. Feature Extraction

Bag of words feature [39] extraction is used in NB, ME and SVM methods, while word2vec creates a feature vector using either Continuous bag of words or Skip gram model which is further used in LSTM and CNN. The methods are explained below.

3.1.1. Bag-of-Words

Bag of words model is a very flexible and simple model used for feature extraction. This model keeps a track of number of occurrences, also called term frequency of every word that appears in the sentence. Also, a specific subjectivity score is assigned to each word of the sentence. The score for each word is added up to find the total score. Depending upon this total score, the polarity of each sentence is decided.

3.1.2. Word2Vec

Word2Vec model is used for forming word embeddings. It is a two-layer neural network created by Tomas Mikolov at google to process text. It takes the text dataset as an input and then outputs a set of vectors [40]. Word2Vec is a combination of two techniques, i.e., Skip-gram model and Continuous bag of words (CBOW) model. This model is very useful as it detects similarities of words in its vector form rather than textual format. These similarities are detected on the basis of word's meaning guessed through its past appearances and association with other words.

3.2. Dictionary-Based Classifier

Valence Aware Dictionary and Sentiment Reasoner [41] (VADER) is a dictionary-based approach that maps words to sentiment by building a or a 'dictionary of sentiment'. In this approach, each word present in the sentence is assigned a score as per the meaning of that word in the dictionary. A final compound score of the sentence is calculated which varies from -1 to 1. This score represents whether the sentence is positive or negative. The compound score for each sentence in the dataset

is combined and an average score for the whole document is analyzed. To compare it with the other machine learning approaches, we convert the average score to accuracy by dividing the score of the whole document by the total number of reviews in that particular data set. VADER focuses on the words used in the sentence and then assigns score to each word based on the dictionary.

3.3. Machine Learning Based Classifiers

We discuss in detail five machine learning based algorithms to determine the sentiment accuracy of the dataset.

3.3.1. Naive Bayes

This is a probabilistic model based on the Bag-of-words module to store only the frequencies of each word and ignore their positioning with respect to each other. By using Bayes Theorem, it estimates the probability that a feature set will belong to a particular predefined label. Naive Bayes classification model [42], based on the distribution of words present in the document or sentence, computes the posterior probability that this document or sentence will belong to a particular class. The probability is based on the distribution and frequency of the words rather than their positioning with respect to each other.

$$P(label|features) = \frac{P(label) * P(features|label)}{P(features)}, \tag{1}$$

where P(label|features) determines the probability that a feature set belongs to a particular label. P(label) is the prior estimate of the label. P(features|label) is the probability that the given feature set belongs to this particular label and P(features) is the prior estimate that this given feature set occurred. However, this classification system makes one fundamental assumption, i.e., words in a reviews, category pair occur independent of other words.

3.3.2. Maximum Entropy

Maximum Entropy (ME) [43] belongs to the class of exponential models. Its polarity is more based on the positioning of words rather than their frequencies. It does not assume that all the features are independent of each other like Naive Bayes. Based on the principle of ME, from all the models, we pick the one that has the largest entropy. The ME classifier uses encoding to convert the feature sets into vectors. Then for computation of most likely labels for each feature set, we combine the calculated weight for each feature [44].

The Maximum Entropy modeling technique provides a probability distribution that is as close to the uniform distribution, so its result is better than Naive Bayes.

3.3.3. Support Vector Machine (SVM)

Support Vector Networks works for multiple machine learning problems such as regression and classification. The main principle that works behind SVM is finding a particular linear classifier that separates all the classes in the search space in the best possible manner. After the pre-processing of the reviews, the improved feature sets were used for sentiment classification, i.e., positive and negative reviews. With the help of hyper plane in support vector machine the data is divided into two classes such as positive and negative. This hyperplane used to map the new examples or the data in the test cases in the same search plane and predict the class to which the data example has more probability of belonging [45].

3.3.4. Long Short Term Memory (LSTM)

Recurrent Neural Networks (RNN) focus on the issue of considering the past information so as to understand the meaning of current and next words. LSTM network [46] is a type of RNN that is

capable of handling long term dependencies as otherwise it was difficult for RNN to connect multiple long term dependencies [47]. After being first introduced by Hochreiter and Schmidhuber in 1997, LSTM has gone through multiple changes over the years. LSTM solves the problem of vanishing and exploding gradient [48], which is a severe limitation for RNN.

The steps of LSTM are defined as: The first step is to decide the information that is going to be deleted from the memory cell. A sigmoid layer executes this decision after looking at prior information i_{t-1} and current input c_t. This sigmoid layer outputs a number between 0 and 1 that determines the amount of information that needs to be retained based on weight W_o. o_t represents the output of the current cell, and b_o is the bias for this particular cell.

$$o_t = \sigma(W_o * [i_{t-1}, c_t] + b_o). \tag{2}$$

Next, it decides the new information that is to be updated into the memory cell. It is done through two steps, a sigmoid layer to decide the values to update and a *tanh* layer to create a vector of new values. n_t denotes the information that is to be updates based on weight W_n and bias b_n and $\tilde{V}_t$ is the data to be included in the current state information. An LSTM cell is shown in Figure 3.
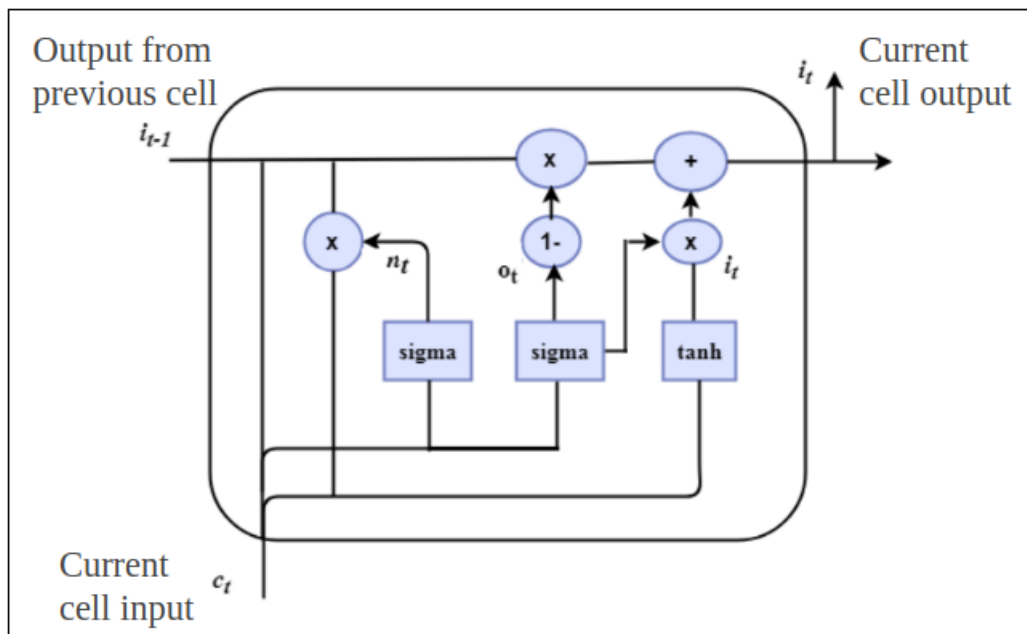

Figure 3. Long Short Term Memory cell, the data flow is from left to right where the current cell input parameter is c_t, i_{t-1} is the output from the previous LSTM cell containing prior information, which is forwarded to the current cell. Both these values are concatenated based on the parameters n_t which denotes the information that is to be updated, o_t which represents the output within the current cell giving the final output value for this layer as i_t that serves as prior information to the next LSTM cell.

$$n_t = \sigma(W_n * [i_{t-1}, c_t] + b_n), \tag{3}$$

$$\tilde{V}_t = tanh(W_V * [i_{t-1}, c_t] + b_V). \tag{4}$$

Now this information is updated into the next cell V_t by multiplying the old state with o_t.

$$V_t = o_t * V_{t-1} + n_t * \tilde{V}_t \tag{5}$$

In the last step, we again implement a sigmoid layer to find f_t that denotes the information which will be given as output based on weight W_f and bias b_f. The *tanh* layer updates the required parts and gives i_t as the output of the cell.

$$f_t = \sigma(W_f * [i_{t-1}, c_t] + b_f), \tag{6}$$

$$i_t = f_t * tanh(V_t). \tag{7}$$

The final output i_t from this cell will serve as prior information for the next cell to find out its subsequent cell state. Nowdays, LSTM are increasingly used to classify test data over other classification algorithms. It is trained on book review dataset with 32 neurons per layer followed by a sigmoid activation function. The netwok has been trained on different epochs and achieved good accuracy compare to other algorithms.

3.3.5. Convolution Neural Network (CNN)

CNN was originally developed for computer vision and its applications, it makes use of local features of the image on which multiple layers with convolving features can be implemented. To implement CNN on the textual reviews, we train a CNN model [49] on book reviews dataset with a single layer on top of the features extracted from the sentences using the word2vec model. First layer is the convolution layer where we slide multiple filters of different sizes over the 128 word embeddings dimensions to produce a feature map based on the particular filter. Max-pooling layer follows this by convolving the results of previous layer into one long feature vector. Max pooling layer finds the most prominent feature vector from the feature map belonging to every filter, which is then passed on to fully connected softmax layer. Dropout regularization is performed before we use softmax layer to classify the result. Regularization randomly drops out some hidden units from the layer to prevent the co-adaptation on training data which may lead to over-fitting. This network is shown in Figure 4.
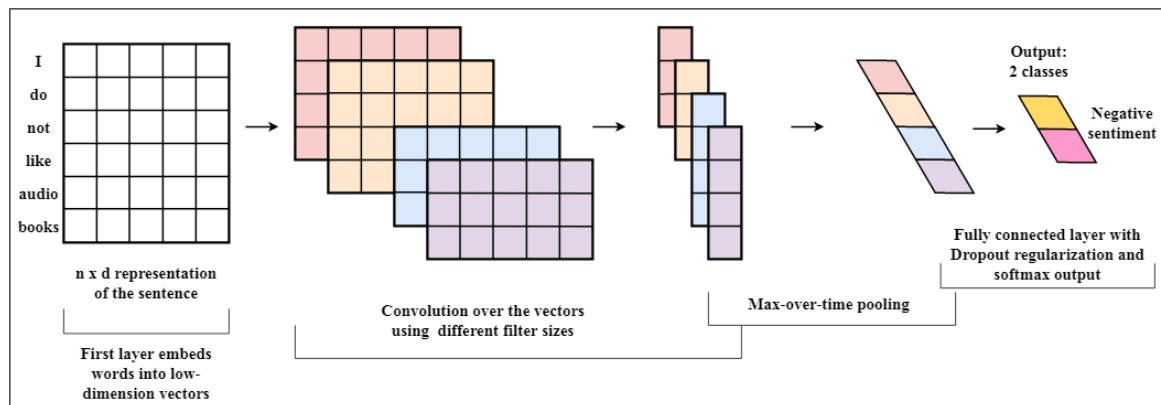

Figure 4. First layers of the model form low-dimensional vectors from the sentence words. The convolution is done by the next layer, using multiple filter sizes such as sliding over 3 or 4 words at a time. Next, the result is max-pooled into a long feature vector and the final results is given using a softmax layer after adding dropout regularization.

4. Experiments and Discussion

In this section, we first describe the dataset, explaining the process of data collection and its further processing that we have done in our experiment. We present the results (see Sections 4.2.1 and 4.2.2) obtained from the feature extraction methods and different classifiers.

4.1. Dataset Description

One of the most crucial parts of this study is data collection. Generally, datasets for sentiment analysis are easily available on the internet which can not be used here as along with the expressed opinion. The micro-blogging and other sites like twitter, Facebook, Amazon, Goodreads, and IMDb do not divulge their user's personal information due to privacy concerns so we create a new dataset that contains all the required information.

The dataset for this experiment is created by collecting opinions of nearly 900 users from the social networking site Facebook. The users have answered a questionnaire containing multiple questions that ask their reviews on preferences of book medium as a Google Form. The questionnaire consisted of questions based on the user's opinions regarding kindles, paperbacks, hardcover, picture, and audiobooks. Further, the questionnaire discusses if the user's thought that digital mediums such as kindle or ebooks could replace hardcover or paperbacks for them. The questions elaborated on whether the user liked audiobooks better than other formats and a short description of their opinions. The form registers the user's opinion, along with the gender and age groups to which they belonged. Along with the user opinions, they have also stated their preference as a positive/negative opinion that serves as the ground truth for the classifiers.

We have selected this domain because we intended to avoid topics with unbalanced spectrum of audience like sports, fashion or television that leaned more towards a particular gender or age group. The responses given by the users to the questionnaire is shown in Figure 5, from the overall reviews we have received, 60% are positive, while the other 40% are negative. From this dataset, we have also segregated the reviews into separate groups, first based on gender, where we have data in a 70% to 30% division to more opinions expressed by the female users. Based on age demographics, the dataset has four age groups into which the users have identified themselves. From the total reviewers, 40% of them belong to the age group of Below 20. The age group 21–34 has nearly 30%, while 20% are in the 35–50 age group. The rest of the users belong to the oldest age group of Above 50.
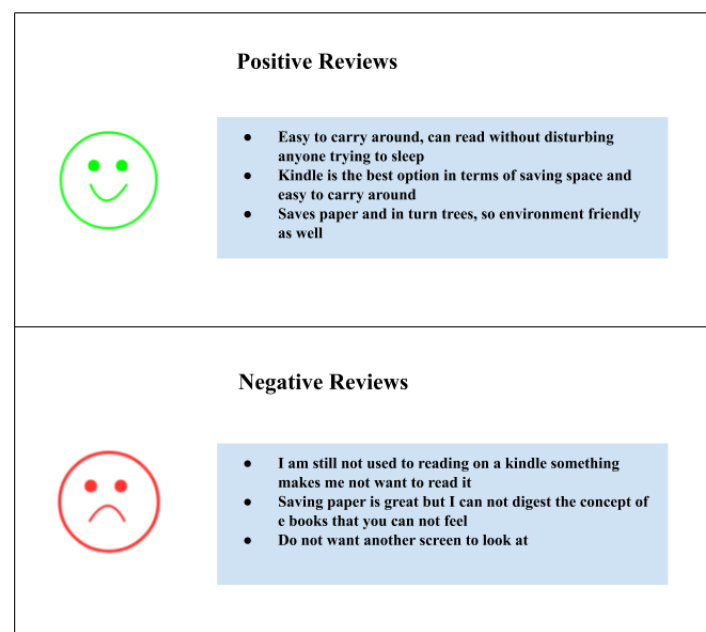

Positive Reviews

- Easy to carry around, can read without disturbing anyone trying to sleep
- Kindle is the best option in terms of saving space and easy to carry around
- Saves paper and in turn trees, so environment friendly as well

Negative Reviews

- I am still not used to reading on a kindle something makes me not want to read it
- Saving paper is great but I can not digest the concept of e books that you can not feel
- Do not want another screen to look at

Figure 5. The collected reviews segregated into positive and negative reviews.

4.2. Result Analysis

We have shown the result of machine learning and dictionary-based approaches on the basis of age and gender information. The results of these classifiers are expressed in terms of accuracy [50].

$$Accuracy = \frac{Correctly\ Predicted\ Observations}{Total\ number\ of\ observations}. \tag{8}$$

4.2.1. Effect of Age

The extracted dataset based on age is divided into four groups: one group with age below 20, second with age from 21 to 34, third from 35 to 50 and the last one with age above 50. Thus, a total four groups are created containing positive and negative responses from people of that particular age group. Another group (without age information) containing reviews from all the age groups is formed to compare its results to the other groups as shown in Figure 6.
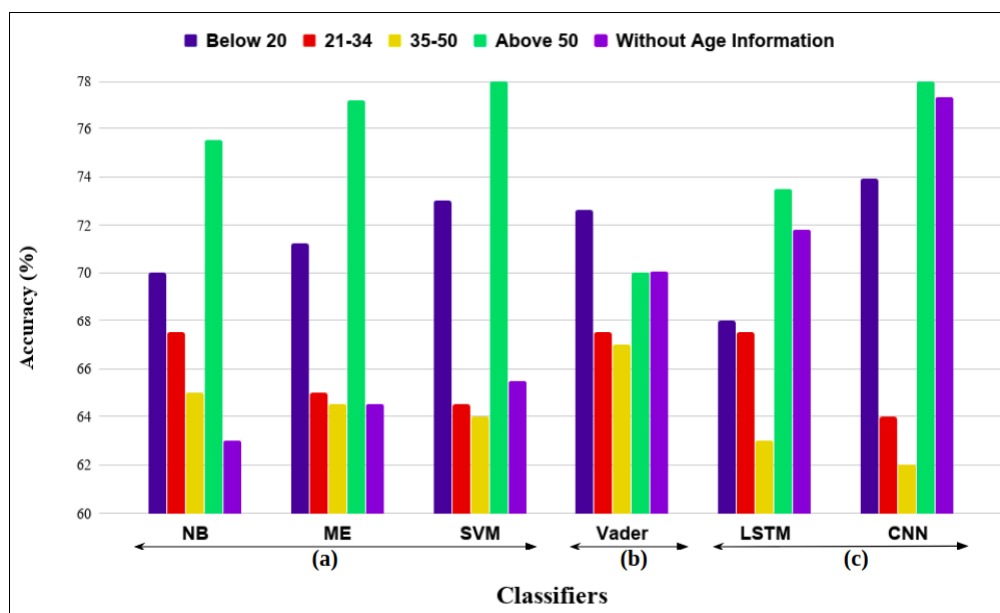

Figure 6. Comparison on basis of Age between different Feature Extraction and Classifier techniques: (**a**) Machine Learning classifiers using Bag-of-words feature extraction method; (**b**) Dictionary-based approach; (**c**) Machine Learning approaches using word2vec feature extraction method.

Pre-processing of all the reviews is performed individually by removing the punctuations, symbols and the stop words from the user reviews as explained in Section 3.1.1. Bag-of-words model on pre-processed data is used to create feature vector which is then used in different classifiers such as NB, ME and SVM. The low dimensional feature vectors are formed from sentences using word2vec model which are then used in LSTM and CNN methods. VADER is also implemented on the pre-processed data. After these approaches are implemented on the separated groups of data individually, the results are recorded.

The 'Above 50' age group performs better as compared to all other age groups in all the classifiers with the highest accuracy of 78% in CNN and SVM classifier. 'Below 20' age group has better accuracy compared to the other two middle age groups where the age group '21–34' performs better than the other age group in all instances, even though the difference between these two age groups are not considerable. Better performance of the eldest age groups shows that the sentiment analysis approaches are able to predict the sentiment in this age group more easily as compared to others groups. The group of data without any age information performs better in LSTM and CNN as compared to other machine learning approaches, where it performs worse than the groups with age information.

4.2.2. Effect of Gender

We label the full dataset into two groups (Male and Female) based on gender containing their positive and negative reviews. Pre-processing, feature extraction and different classifiers are implemented on these data groups similarly as in Section 3. The results are represented in Figure 7.
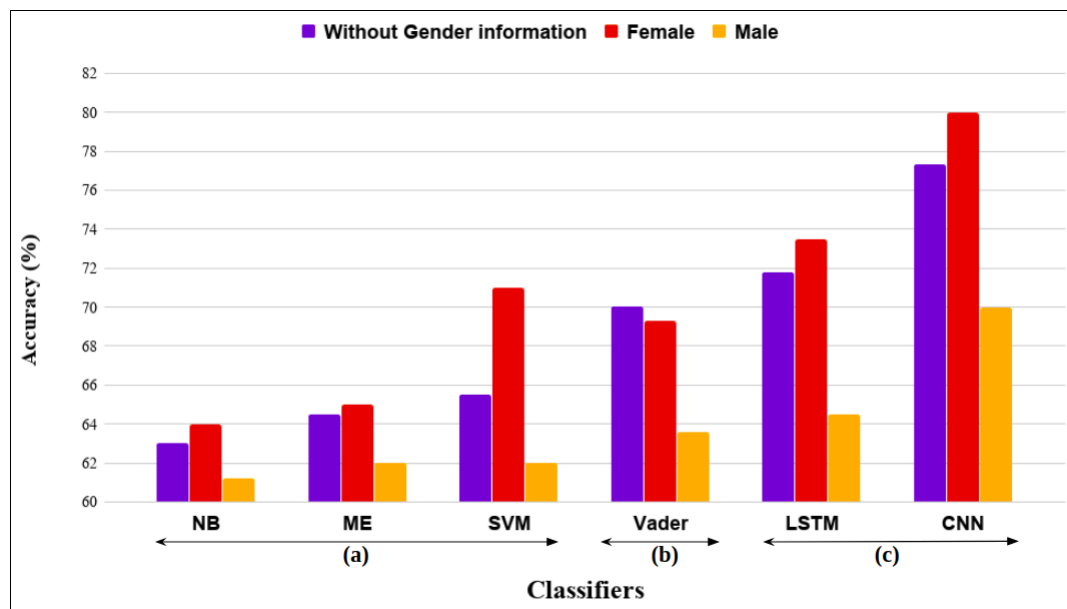

Figure 7. Comparison on basis of Gender between different Feature Extraction and Classifier techniques: (**a**) Machine Learning classifiers using Bag-of-words feature extraction method; (**b**) Dictionary-based approach; (**c**) Machine Learning approaches using word2vec feature extraction method.

It can be clearly seen that female data generates better accuracy as compared to the data without gender information and the male data. Female data has the best accuracy in CNN classifier of 80%, which is better than the other classifiers. This result aligns with the psychological studies that females express their opinion better as compared to their male counterparts. The sentiment in female data is easier to predict, hence giving a better accuracy. This pattern of female data having better accuracy can be observed in all the machine learning approaches.

5. Conclusions and Future Work

In this paper, we have compared multiple sentiment analysis techniques on the dataset collected from nearly 900 users from Facebook along with the users' age and gender information. We extracted this dataset into four groups to analyze the impact of age and gender on the way the user expresses his/her opinion. Machine learning and Dictionary-based techinques have been performed to know the sentiment analysis of the reviews. With respect to gender, female data recorded the best accuracy while for age, the Above 'Age 50' group has the better accuracy as compared to all other age groups. The results can be further improved by collecting more data for both male and female and different age groups.

In future work, we can also include exploration of reviews in audio and visual format to detect emotions from the way of speech and facial expressions of the user to provide more comprehensive investigations from different aspects.

Author Contributions: All authors have contributed to this paper. M.G. and S.K. proposed the main idea, worked on the introduction and data collection. M.G., S.K. and P.P.R. were involved in the methodology and M.G. performed the analyses. M.G. and S.K. drafted the manuscript. B.-G.K., P.P.R. and D.P.D. contributed to the final version of the paper. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declared that they have no conflicts of interest to this work.

References

1.　Manek, A.S.; Shenoy, P.D.; Mohan, M.C.; Venugopal, K. Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier. *Worldw. Web* **2017**, *20*, 135–154. [CrossRef]

2.　Dos Santos, C.; Gatti, M. Deep convolutional neural networks for sentiment analysis of short texts. In Proceedings of the COLING, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 23–29 August 2014; pp. 69–78.

3.　Kiritchenko, S.; Zhu, X.; Cherry, C.; Mohammad, S. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews. In Proceedings of the 8th International Workshop on Semantic Evaluation, Dublin, Ireland, 23–24 August 2014; pp. 437–442.

4.　Pontiki, M.; Galanis, D.; Papageorgiou, H.; Manandhar, S.; Androutsopoulos, I. Semeval-2015 task 12: Aspect based sentiment analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation, Denver, CO, USA, 4–5 June 2015; pp. 486–495.

5.　Cao, D.; Ji, R.; Lin, D.; Li, S. A cross-media public sentiment analysis system for microblog. *Multim. Syst.* **2016**, *22*, 479–486. [CrossRef]

6.　Ghosh, R.; Zhang, L.; Dekhil, M.E.; Liu, B. Performing sentiment analysis on microblogging data, including identifying a new opinion term therein. US Patent 9,275,041, 1 March 2016.

7.　Ullah, M.A.; Islam, M.M.; Azman, N.B.; Zaki, Z.M. An overview of multimodal sentiment analysis research: Opportunities and difficulties. In Proceedings of the 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition, Himeji, Japan, 1–3 September 2017; pp. 1–6.

8.　Cambria, E. Affective computing and sentiment analysis. *IEEE Intell. Syst.* **2016**, *31*, 102–107. [CrossRef]

9.　Liu, B. Sentiment analysis and subjectivity. *Handb. Nat. Lang. Proc.* **2010**, *2*, 627–666.

10.　Kumar, S.; Yadava, M.; Roy, P.P. Fusion of eeg response and sentiment analysis of products review to predict customer satisfaction. *Inf. Fus.* **2019**, *52*, 41–52. [CrossRef]

11.　Kim, J.H.; Kim, B.G.; Roy, P.P.; Jeong, D.M. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. *IEEE Access* **2019**, *7*, 41273–41285. [CrossRef]

12.　Yoo, S.M.; Cho, C.; Lee, K.H.; Park, J.; Jin, S.; Lee, Y.; Kim, B.G. Structure of deep learning inference engines for embedded systems. In Proceedings of the IEEE 2019 International Conference on Information and Communication Technology Convergence, Kuala Lumpur, Malaysia, 24–26 July 2019; pp. 920–922.

13.　Kim, J.H.; Hong, G.S.; Kim, B.G.; Dogra, D.P. Deepgesture: Deep learning-based gesture recognition scheme using motion sensors. *Displays* **2018**, *55*, 38–45. [CrossRef]

14.　Kahaki, S.M.M.; Ismail, W.; Nordin, M.J.; Ahmad, N.S.; Ahmad, M. Automated age estimation based on geometric mean projection transform using orthopantomographs. *J. Adv. Technol. Eng. Stud.* **2017**, *3*, 6–10.

15.　Kahaki, S.M.; Nordin, M.J.; Ahmad, N.S.; Arzoky, M.; Ismail, W. Deep convolutional neural network designed for age assessment based on orthopantomography data. *Neural Comput. Appl.* **2019**, *3*, 1–12. [CrossRef]

16.　Li, Y.M.; Li, T.Y. Deriving market intelligence from microblogs. *Decis. Support Syst.* **2013**, *55*, 206–217. [CrossRef]

17.　Lockenhoff, C.E.; Costa, P.T.; Lane, R.D. Age differences in descriptions of emotional experiences in oneself and others. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **2008**, *63*, 92–99. [CrossRef] [PubMed]

18.　Zimmermann, P.; Iwanski, A. Emotion regulation from early adolescence to emerging adulthood and middle adulthood. *Int. J. Behav. Dev.* **2014**, *38*, 182–194. [CrossRef]

19.　Kaur, B.; Singh, D.; Roy, P.P. Age and gender classification using brain–computer interface. *Neural Comput. Appl.* **2019**, *31*, 5887–5900. [CrossRef]

20. Oh, H.; Parks, S.C.; Demicco, F.J. Age-and gender-based market segmentation: A structural understanding. *Int. J. Hosp. Tour. Adm.* **2002**, *3*, 1–20. [CrossRef]

21. Keshari, P.; Jain, S. Effect of age and gender on consumer response to advertising appeals. *Paradigm* **2016**, *20*, 69–82. [CrossRef]

22. Appel, O.; Chiclana, F.; Carter, J.; Fujita, H. Successes and challenges in developing a hybrid approach to sentiment analysis. *Appl. Intell.* **2018**, *48*, 1176–1188. [CrossRef]

23. Zainuddin, N.; Selamat, A.; Ibrahim, R. Hybrid sentiment classification on twitter aspect-based sentiment analysis. *Appl. Intell.* **2018**, *48*, 1218–1232. [CrossRef]

24. Fan, T.K.; Chang, C.H. Sentiment-oriented contextual advertising. *Knowl. Inf. Syst.* **2010**, *23*, 321–344. [CrossRef]

25. Kuo, Y.H.; Fu, M.H.; Tsai, W.H.; Lee, K.R.; Chen, L.Y. Integrated microblog sentiment analysis from users' social interaction patterns and textual opinions. *Appl. Intell.* **2016**, *44*, 399–413. [CrossRef]

26. Li, G.; Liu, F. Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions. *Appl. Intell.* **2014**, *40*, 441–452. [CrossRef]

27. Ali, F.; Kim, E.K.; Kim, Y.G. Type-2 fuzzy ontology-based opinion mining and information extraction: A proposal to automate the hotel reservation system. *Appl. Intell.* **2015**, *42*, 481–500. [CrossRef]

28. Pak, A.; Paroubek, P. Twitter as a corpus for sentiment analysis and opinion mining. In *LREC*; University of Paris: Paris, France, 2010; pp. 1320–1326.

29. Denecke, K. Using sentiwordnet for multilingual sentiment analysis. In Proceedings of the IEEE 24th International Conference on Data Engineering Workshop, Cancun, Mexico, 7–12 April 2008; pp. 507–512.

30. Zhou, G.; Zhu, Z.; He, T.; Hu, X.T. Cross-lingual sentiment classification with stacked autoencoders. *Knowl. Inf. Syst.* **2016**, *47*, 27–44. [CrossRef]

31. Cheng, J.; Zhang, X.; Li, P.; Zhang, S.; Ding, Z.; Wang, H. Exploring sentiment parsing of microblogging texts for opinion polling on chinese public figures. *Appl. Intell.* **2016**, *45*, 429–442. [CrossRef]

32. Fabes, R.A.; Martin, C.L. Gender and age stereotypes of emotionality. *Personal. Soc. Psychol. Bull.* **1991**, *17*, 532–540. [CrossRef]

33. Stoner, S.B.; Spencer, W.B. Age and gender differences with the anger expression scale. *Educ. Psychol. Meas.* **1987**, *47*, 487–492. [CrossRef]

34. Davis, T.L. Gender differences in masking negative emotions: Ability or motivation? *Dev. Psychol.* **1995**, *31*, 660–667. [CrossRef]

35. Brody, L.R. Gender and emotion: Beyond stereotypes. *J. Soc. Issues* **2010**, *53*, 369–393. [CrossRef]

36. Kring, A.M.; Gordon, A.H. Sex differences in emotion: Expression, experience, and physiology. *J. Personal. Soc. Psychol.* **1998**, *74*, 686–703 . [CrossRef]

37. Birditt, K.S.; Fingerman, K.L. Age and gender differences in adults' descriptions of emotional reactions to interpersonal problems. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **2003**, *58*, 237–245. [CrossRef]

38. Kharde, V.A.; Sonawane, S.S. Sentiment analysis of Twitter data: A survey of techniques. *Int. J. Comput. Appl.* **2016**, *139*, 5–15.

39. Saini, R.; Kumar, P.; Roy, P.P.; Pal, U. Trajectory classification using feature selection by genetic algorithm. In *Proceedings of the 3rd International Conference on Computer Vision and Image Processing*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 377–388.

40. Xue, B.; Fu, C.; Shaobin, Z. A study on sentiment computing and classification of Sina Weibo with Word2vec. In Proceedings of the IEEE International Congress on Big Data, Anchorage, AK, USA, 27–July 2 June 2014; pp. 358–363.

41. Hutto, C.J.; Gilbert, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the 8th International Conference on Weblogs and Social Media, Ann Arbor, MI, USA, 1–4 June 2014; pp. 216–255.

42. Zhang, H. The optimality of naive bayes. *AA* **2004**, *1*, 1–6.

43. Nigam, K.; Lafferty, J.; McCallum, A. Using maximum entropy for text classification. In *IJCAI-99 Workshop on Machine Learning for Information Filtering*; IJCAI: Stockholom, Sweden, 1999; Volume 1, pp. 61–67.

44. Kaufmann, J.M. JMaxAlign: A maximum entropy parallel sentence alignment tool. In Proceedings of the COLING 2012: Demonstration Papers, Mumbai, India, 8–15 December 2012; pp. 277–288.

45. Mullen, T.; Collier, N. Sentiment analysis using support vector machines with diverse information sources. In Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, Barcelona, Spain, 25–26 July 2004; pp. 412–418.

46.   Saini, R.; Kumar, P.; Kaur, B.; Roy, P.P.; Dogra, D.P.; Santosh, K. Kinect sensor-based interaction monitoring system using the blstm neural network in healthcare. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 2529–2540. [CrossRef]

47.   Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

48.   LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]

49.   Kim, Y. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1746–1751.

50.   Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE* **2015**, *10*, e0118432. [CrossRef] [PubMed]